*Proceedings of the 2001 Winter Simulation Conference*
*B. A. Peters, J. S. Smith, D. J. Medeiros, and M. W. Rohrer, eds.*

# USING COMMON RANDOM NUMBERS FOR INDIFFERENCE-ZONE SELECTION

E. Jack Chen

Mathematical Modeling Group
BASF Corporation
3000 Continental Drive
Mount Olive, New Jersey 07828, U.S.A.

## ABSTRACT

This paper discusses the validity of using common random numbers (CRNs) with two-stage selection procedures to improve the possibility of correct selection and discusses the intrinsic subset pre-selection of the *Enhanced Two-Stage Selection* (ETSS) procedure. We propose using CRNs with *Rinott's two-stage selection* and ETSS procedures when the underlying processes satisfy certain conditions. An experimental performance evaluation demonstrates the improvement in the possibility of correct selection of using CRNs with two-stage selection procedures and the intrinsic subset pre-selection of the ETSS procedure.

## 1 INTRODUCTION

Discrete-event simulation has been widely used to compare alternative system designs or operating policies. When evaluating $k$ alternatives, we often would like to select one as the best and to control the probability that the selected alternative really is best. Let $\mu_i$ denote the (unknown) expected response of alternative $i$. Our goal is to find the alternative with the smallest expected response $\mu^* = \min_{1 \le i \le k} \mu_i$. If the goal is to select an alternative with the biggest expected response, just replace min by max in expression for $\mu^*$. We achieve this goal by using a class of ranking and selection (R&S) procedures.

Many R&S procedures are directly or indirectly based on Dudewicz and Dalal's (1975) or Rinott's (1978) indifference-zone-selection procedures. However, their procedures require the simulation output sequence is independent and normally distributed. It is known that using common random numbers (CRNs) to the simulation of each design can reduce the variance of estimates of the pairwise difference of sample means. However, because the assumption of independent samples across designs is used to develop the two-stage selection procedures, CRNs are generally not used with those procedures for variance reduction. For an overview of existing methods of R&S see Bechhofer, Santner, and Goldsman (1995), Goldsman and Nelson (1998), or Law and Kelton (2000).

Nelson and Matejcik (1995) develop a selection procedure that does not require independent samples across systems. They compare their procedure with Clark and Yang's (1986) procedure, which is based on the Bonferroni inequality. Both procedures are valid with CRNs. Nelson and Matejcik (1995) point out that procedures based on the Bonferroni inequality become more conservative as the number of alternatives, $k$, increases. At some point this conservatism overwhelms the benefit from CRNs. Instead, their procedure is based on the assumption that the covariance matrix has a particular structure known as *sphericity*. Chick and Inoue (2000) compare the performance of the indifference-zone procedures of Nelson and Matejcik (1995), Clark and Yang (1986) and their Bayesian-based procedures. Both indifference-zone procedures are a bit more difficult to apply than Dudewicz and Dalal or Rinott's procedure because the first-stage data from all $k$ systems must be saved to compute the required sample size at the second-stage.

Let CS denote the event of "correct selection." In a stochastic simulation such a CS can never be guaranteed with certainty. The possibility of CS denoted by $P(CS)$, is a random variable dependent on sample sizes and other uncontrollable factors. Let $P_I(CS)$ denote the possibility of correct selection when samples are independent within each alternative and across alternatives, and $P_C(CS)$ denote the possibility of correct selection when CRNs are used, i.e., independent sampling within each alternative but not across alternatives. The positive effects of using CRNs with Rinott and ETSS procedures is not guaranteed. However, our experimental results show that under certain conditions when input random numbers are positively correlated (a necessary condition for CRNs to be effective), the possibility of correct selection $P_C(CS)$ is larger than $P_I(CS)$ for Rinott and ETSS (Chen and Kelton 2000c) procedures.

In Section 2 we provide background necessary to understand our proposed procedure. In Section 3 we present our proposed procedure for selection. In Section 4, we

show our empirical-experiment results. In Section 5, we give concluding remarks.

## 2 BACKGROUND

This section gives some general background of ranking and selection procedures and does not involve CRNs. The use of CRNs will be introduced in Section 3.2.

First, some notation:

- $X_{ij}$: the average of the observations from the $j^{th}$ replication or batch of the $i^{th}$ alternative,
- $N_i$: the number of replications or batches for alternative $i$,
- $\mu_i$: the expected performance measure for alternative $i$, i.e., $\mu_i = E(X_{ij})$,
- $\hat{\mu}_i$: the sample mean performance measure for alternative $i$, i.e., $\sum_{j=1}^{N_i} X_{ij}/N_i$,
- $\sigma_i^2$: the variance of the observed performance measure of alternative $i$ from one replication or batch, i.e., $\sigma_i^2 = \mathrm{Var}(X_{ij})$,
- $S_i^2(N_i)$: the sample variance of alternative $i$ with $N_i$ replications or batches, i.e., $S_i^2(N_i) = \sum_{j=1}^{N_i}(X_{ij} - \hat{\mu}_i)^2/(N_i - 1)$.

### 2.1 Indifference-Zone-Selection Procedures

Let $\mu_{i_l}$ be the $l^{th}$ smallest of the $\mu_i$'s, so that $\mu_{i_1} \leq \mu_{i_2} \leq \ldots \leq \mu_{i_k}$. Our goal is to select an alternative with the smallest expected response $\mu_{i_1}$. In practice however, if $\mu_{i_1}$ and $\mu_{i_2}$ are very close together, we might not care if we mistakenly choose alternative $i_2$, whose expected response is $\mu_{i_2}$. The "practically significant" difference $d^*$ (a positive real number) between a best and next-best alternative is called the indifference zone in the statistical literature and represents the smallest difference about which we care. Therefore, we want a procedure that avoids making a large number of replications or batches to resolve differences less than $d^*$. That is, we want $P(CS) \geq P^*$ provided that $\mu_{i_2} - \mu_{i_1} \geq d^*$, where the minimal CS probability $P^*$ and the "indifference" amount $d^*$ are both specified by the user.

### 2.2 The Two-Stage Rinott Procedure

The two-stage procedure of Rinott (1978) has been widely studied and applied. Let $n_0$ be the number of initial replications or batches from each alternative. The first-stage sample means $\hat{\mu}_i^{(1)} = \sum_{j=1}^{n_0} X_{ij}/n_0$, and sample variances

$$S_i^2(n_0) = \frac{\sum_{j=1}^{n_0}(X_{ij} - \hat{\mu}_i^{(1)})^2}{n_0 - 1},$$

for $i = 1, 2, \ldots, k$ are computed. Based on the number of initial replications or batches $n_0$ and the sample variances $S_i^2(n_0)$ obtained from the first stage, the number of additional simulation replications or batches for each alternative in the second stage is $N_i - n_0$, where

$$N_i = \max(n_0, \lceil (hS_i(n_0)/d^*)^2 \rceil), \text{ for } i = 1, 2, \ldots, k. \quad (1)$$

Here $\lceil z \rceil$ is the smallest integer that is greater than or equal to the real number $z$, and $h$ (which depends on $k$, $P^*$, and $n_0$) is a constant which solves Rinott's (1978) integral ($h$ can also be found from tables in Wilcox, 1984). We then compute the overall sample means $\hat{\mu}_i = \sum_{j=1}^{N_i} X_{ij}/N_i$, and select the alternative with the smallest $\hat{\mu}_i$ as best. Basically, the computing budget is allocated proportional to the sample variances from the first stage.

### 2.3 Subset Pre-Selection

Goldsman et al. (1999) describe a procedure to pre-select a subset of alternatives, excluding the other (inferior) alternatives from further consideration. This procedure is denoted as SSPS in our discussion.

Let $\hat{\mu}_{j_l}$ be the $l^{th}$ smallest of the $\hat{\mu}_i$'s from the first stage, so that $\hat{\mu}_{j_1} \leq \hat{\mu}_{j_2} \leq \ldots \leq \hat{\mu}_{j_k}$. Then $\hat{\mu}_{j_1}$ is automatically included in the subset. For all $i \neq j_1$ , let $t_{1-\alpha,n}$ be the $1-\alpha$ quantile of the $t$ distribution with $n$ degrees of freedom (df)

$$t = t_{1-(P^*)^{\frac{1}{k-1}}, n_0-1},$$

$$S_{i,j_1}^2 = \frac{\sum_{j=1}^{n_0}(X_{ij} - X_{j_1 j} - (\hat{\mu}_i - \hat{\mu}_{j_1}))^2}{n_0 - 1}, \quad (2)$$

and

$$W_{i,j_1} = tS_{i,j_1}/\sqrt{n_0}.$$

Then alternative $i$ will be included in the subset only if $\hat{\mu}_i - \hat{\mu}_{j_1} \leq (W_{i,j_1} - d^*)^+$, where $(a)^+ = \max(0, a)$.

### 2.4 An Enhanced Two-Stage Selection (ETSS) Procedure

Chen and Kelton (2000c) propose an ETSS procedure that takes into account not only the sample variances but also the difference of sample means across alternatives. They show that ETSS procedure achieves $P(CS) \geq P^*$ and significantly reduces the number of total simulation replications or batches when compares to Rinott's procedure. Let

$$r_i = \max(\hat{\mu}_i - \hat{\mu}_{j_1}, d^*)/d^*, \quad (3)$$

and

$$h_i = h/r_i, \quad (4)$$

(here $h$ has the same value as Rinott's procedure) then ETSS procedure computes the number of required simulation replications or batches for each alternative based on the following formula

$$N_i = \max(n_0, \lceil (h_i S_i(n_0)/d^*)^2 \rceil), \text{ for } i = 1, 2, \ldots, k. \quad (5)$$

The difference between equations (5) and (1) is that $h_i$ instead of $h$ is used. The information of the difference of the sample means between alternative designs is embedded in the value of $h_i$ through $r_i$ of equation (3), thus, equation (5) utilizes the information of both the sample means and variances.

## 3 METHODOLOGIES

In this section we show the validity of using CRNs with Rinott and ETSS procedures to increase $P(CS)$. If the input data are not i.i.d. normal, users can use batch means (see Chen and Kelton 2000a, b) to obtain sample means that are essentially i.i.d. normal.

### 3.1 Prologue

Rinott's procedure assumes input data are i.i.d. normal and is based on the *least favorable configuration* (LFC), i.e., $\mu_{i_1} + d^* = \mu_{i_2} = \ldots = \mu_{i_k}$. Without loss of generality, assuming $\mu_{i_l} - \mu_{i_1} \geq d^*$ for $i_l \neq i_1$. The possibility of correct selection is

$$P_I(CS) = P(\hat{\mu}_{i_1} < \hat{\mu}_{i_l}, \text{ for } l = 2, 3, \ldots, k).$$

Let

$$T_i = \frac{\hat{\mu}_i - \mu_i}{d^*/h},$$

then

$$P_I(CS) = P(T_{i_l} \geq T_{i_1} - \frac{\mu_{i_l} - \mu_{i_1}}{d^*/h}, \text{ for } l = 2, 3, \ldots, k).$$

If we let $f$ and $F$ denote the density and distribution function, respectively, of the $t$ distribution with $n_0 - 1$ degrees of freedom. By conditioning on $T_{i_1} = t$ and based on the LFC, it can be shown that

$$P_I(CS) \geq \int_{-\infty}^{\infty} F(h + t)^{k-1} f(t) dt \geq P^* \quad (6)$$

see Rinott (1978).

### 3.2 Using Common Random Numbers

Rinott's procedure is derived on the assumption that samples across alternatives are independent. We show that using CRNs can improve the probability of correct selection for a given sample size. We assume that, under properly synchronized CRNs, $Cov[X_{aj}, X_{bj}] > 0$, furthermore, $Cov[X_{aj}, X_{a'j'}] = 0$, for any $a, a'$ and $j \neq j'$, because samples are independent within each alternative. Thus, $Cov[\hat{\mu}_a, \hat{\mu}_b] > 0$ for $a \neq b$. Therefore,

$$Var[\hat{\mu}_a - \hat{\mu}_b] = Var[\hat{\mu}_a] + Var[\hat{\mu}_b] - 2Cov[\hat{\mu}_a, \hat{\mu}_b] \quad (7)$$

will be smaller than in the case of independent sampling. While there are examples of CRNs "backfiring" and causing $Cov[\hat{\mu}_a, \hat{\mu}_b] < 0$ (e.g., Wright and Ramsay 1979), most analyst agree that such instances are rare and we do not consider such situations. We will start our discussion with any pair of alternatives $\mu_{i_1}$ and $\mu_{i_l}$ for $l = 2, \ldots, k$. Thus, with CRNs

$$P_C(CS) = P[\hat{\mu}_{i_1} < \hat{\mu}_{i_l} | \mu_{i_1} < \mu_{i_l}].$$

The standard deviation within each alternative from the first stage is not influenced by CRNs. Thus, the required number of simulation replications or batches for each alternative determined by equations (1) or (5) should be consistent regardless of whether CRNs are used. However, since $Var[\hat{\mu}_{i_1} - \hat{\mu}_{i_2}]$ should be smaller when CRNs are used, it is less likely that $\hat{\mu}_{i_1} - \hat{\mu}_{i_l}$ will be positive, i.e., that CS will not occur, since $\mu_{i_1} - \mu_{i_l} < 0$.

**Proposition I:** For any pairwise comparison between $\hat{\mu}_{i_1}$ and $\hat{\mu}_{i_l}$, the probability of correct selection of two-stage selection procedures is improved when we use CRNs, i.e., $P_C(\hat{\mu}_{i_1} < \hat{\mu}_{i_l}) \geq P_I(\hat{\mu}_{i_1} < \hat{\mu}_{i_l})$. **Proof:** See Appendix A.

Let random variable $Y_l = \hat{\mu}_{i_l} - \hat{\mu}_{i_1}$ and event $E_l$ denote $Y_l > 0$, for $l = 2, 3, \ldots, k$. We can generalize the $P(CS)$ when the number of alternative is $k \geq 2$.

$$\begin{aligned} & P(CS) \\ = \; & P(E_2 \text{ and } E_3 \text{ and } \ldots \text{ and } E_k) \\ = \; & P(E_2) P(E_3|E_2) \ldots P(E_k|E_2, E_3, \ldots, E_{k-1}). \end{aligned}$$

The covariance of $Y_i$ and $Y_j$ is

$$\begin{aligned} & Cov(Y_i, Y_j) \\ = \; & Cov(\hat{\mu}_i - \hat{\mu}_{i_1}, \hat{\mu}_j - \hat{\mu}_{i_1}) \\ = \; & Cov(\hat{\mu}_i, \hat{\mu}_j) - Cov(\hat{\mu}_i, \hat{\mu}_{i_1}) - \\ & Cov(\hat{\mu}_{i_1}, \hat{\mu}_j) + Cov(\hat{\mu}_{i_1}, \hat{\mu}_{i_1}). \end{aligned}$$

In the case of independent sampling across alternatives, $Cov(\hat{\mu}_i, \hat{\mu}_j) = 0$, for $i \neq j$. Therefore, $Cov(Y_i, Y_j) = Var(\hat{\mu}_{i_1}) > 0$ and the correlation coeffi-

cient $\rho_I = \frac{Cov(Y_i,Y_j)}{\sigma_{Y_i}\sigma_{Y_j}} > 0$, i.e., $E_i$ and $E_j$ are positively correlated. Thus, for $l = 3, \ldots, k$,

$$P_I(E_l|E_2, E_3, \ldots, E_{l-1}) \neq P_I(E_l),$$

and by Slepian's inequality (Tong 1980)

$$P_I(CS) \geq P_I(E_2)P_I(E_3)\ldots P_I(E_k).$$

In the case of CRNs, $Cov(\hat{\mu}_i, \hat{\mu}_j) \geq 0$, for $i \neq j$. Therefore, $Cov(Y_i, Y_j)$ could be positive, zero, or negative. Let consider the case $k = 3$ and assume that $Y_3$ and $Y_2$ are bivariate normal, then the conditional probability density function of $Y_3$ given $Y_2$ $f_{Y_3|Y_2}(y_3, y_2)$ has a normal distribution with mean $\mu_{Y_3} + \rho(y_2 - \mu_{Y_2})\sigma_{Y_3}/\sigma_{Y_2}$ and variance $\sigma_Y^2(1-\rho^2)$. The marginal variance $\sigma_{Y_3}^2$ and $\sigma_{Y_2}^2$ should be the same regardless whether CRNs are used. However, the correlation coefficient with CRNs $\rho_C$ could be positive, zero, or negative. If $\rho_C \geq \rho_I > 0$, then $E_C(Y_3|Y_2 > \mu_{Y_2}) \geq E_I(Y_3|Y_2 > \mu_{Y_2})$, $\sigma_Y^2(1-\rho_C^2) \leq \sigma_Y^2(1-\rho_I^2)$, and $P_C(E_3|E_2) \geq P_I(E_3|E_2)$. Therefore, when we use CRNs with two-stage selection procedures we benefit from $P_C(E_2) \geq P_I(E_2)$ and maybe $P_C(E_l|E_2, E_3, \ldots, E_{l-1}) \geq P_I(E_l|E_2, E_3, \ldots, E_{l-1})$, for $l = 3, 4, \ldots, k$. However, if $\rho_I > \rho_C$, then $E_I(Y_3|Y_2 > \mu_{Y_2}) > E_C(Y_3|Y_2 > \mu_{Y_2})$ and we may have $P_I(E_3|E_2) > P_C(E_3|E_2)$. Thus, using CRNs may "backfire" and causing $P_I(CS) > P_C(CS)$.

The sample size of the ETSS is determined by both the sample variance within each alternative and the difference of sample means across alternatives. If we consider the factor of the sample variance within each alternative alone, we will not only increase $P(CS)$ but also reduce the total number of replications or batches when CRNs are used with the ETSS procedure. This is because $h_i$ is squared in equation (5) and the variance of $\hat{\mu}_i$ is reduced when CRNs are used.

The number of replications or batches allocated in the second stage will be greater (say by $x$ replications or batches) if the sample mean is underestimated ($\hat{\mu}_i = \mu_i - \epsilon$ for some $\epsilon > 0$) than using the true mean $\mu_i$. The number of replications or batches allocated in the second stage will be smaller (say by $y$ replications or batches) if the sample mean is overestimated ($\hat{\mu}_i = \mu_i + \epsilon$) than using the true mean $\mu_i$. Furthermore, $x > y$ and the difference between $x$ and $y$ increases as the amount of the deviation $\epsilon$ from the true mean increases.

For example, let $n_i$ be the number of replications or batches for alternative $i$ at the second stage when we know the true mean $\mu_i$, $n_{i_u}$ be the number of replications or batches for alternative $i$ when $\hat{\mu}_i = \mu_i - \epsilon$, where $\epsilon > 0$ and $n_{i_o}$ be the number of replications or batches for alternative $i$ when $\hat{\mu}_i = \mu_i + \epsilon$. Then $x\ (= n_{i_u} - n_i) > y\ (= n_i - n_{i_o})$ because $(\frac{h}{r_i - r_\epsilon})^2 - (h/r_i)^2 > (h/r_i)^2 - (\frac{h}{r_i + r_\epsilon})^2$, where $r_\epsilon = \epsilon/d^*$ and assuming $\hat{\mu}_i - \hat{\mu}_{j_1} > d^*$. The difference between $n_{i_u} - n_i$ and $n_i - n_{i_o}$ increases as the deviation $\epsilon$ from the true mean increases. Therefore, $\text{ETSS}_C$ (using CRNs with the ETSS procedure) can reduce the number of replications or batches when we reduce the variance of $\hat{\mu}_{i_l}$.

If we consider the factor of the difference of the sample mean $\delta_i = \hat{\mu}_i - \hat{\mu}_{j_1}$, using CRNs with ETSS may increase the total number of replications or batches. This is because with independent sampling when we have unusual large sample mean $\hat{\mu}_{i_1}$ in the first stage $\hat{\mu}_{i_1} > \hat{\mu}_{j_1}$, $\delta_i > \hat{\mu}_i - \hat{\mu}_{i_1}$ and results in smaller total number of replications or batches. With CRNs when we have unusual large sample mean $\hat{\mu}_{i_1}$ in the first stage, sample means of all other alternatives will tend to large as well. Therefore, most probably $\hat{\mu}_{j_1} = \hat{\mu}_{i_1}$ and smaller $\delta_i$ is used and results in larger total number of replications or batches. Because ETSS uses the value $\delta_i$ to compute the number of total replications or batches, using ETSS with CRNs has an additional benefit that it increases the precision of the difference of the sample means $\delta_i$.

Note that the subset selection of SSPS requires independent sampling across alternatives. Therefore, CRNs and SSPS cannot be used at the same time. However, we can use independent sampling at the first stage with subset selection and use CRNs at the second stage of the Rinott and ETSS procedures. Moreover, ETSS will allocate few or no replications or batches for inferior alternatives at the second stage. That is, ETSS has an intrinsic subset pre-selection built-in. Alternative $i$ having $N_i = n_0$ can be viewed as being excluded from further consideration. Furthermore, if $N_{j_1} = n_0$ and $\hat{\mu}_{j_1} = \min_{1 \leq i \leq k} \hat{\mu}_i$, then alternative $j_1$ will still be returned from our procedure. Therefore, we recommend using CRNs with ETSS and bypassing SSPS.

## 4 EMPIRICAL EXPERIMENTS

In this section we present some empirical results with the Rinott, and ETSS procedures under both independent sampling across alternatives (denoted as $\text{Rinott}_I$ and $\text{ETSS}_I$), as well as using CRNs (denoted as $\text{Rinott}_C$ and $\text{ETSS}_C$). We also compare the SSPS and the intrinsic pre-selection of the ETSS procedure. The purpose of the experiments was not only to test the methods thoroughly, but also to demonstrate empirically the efficiency improvement of using CRNs with these two procedures. For the first three experiments, the CRNs are completely synchronized, therefore, the covariance structures should be as discussed in the beginning of Section 3.2. On the other hand, the CRNs are not completely synchronized in experiment 4, because a portion of the same random number streams may be used to generate $X_{aj}$ as well as $X_{a'j'}$, where $a \neq a'$ and $j \neq j'$. Therefore, the covariance structures may deviate from our assumptions, i.e., for some $a \neq a'$ and $j \neq j'$ $Cov[X_{aj}, X_{a'j'}] > 0$.

Table 1: P(CS) and Sample Sizes for Experiment 1

| | | $P^* = 0.90$ | | $P^* = 0.95$ | |
|---|---|---|---|---|---|
| $n_0$ | Procedure | $P(CS)$ | $\overline{T}$ | $P(CS)$ | $\overline{T}$ |
| 20 | Rinott$_C$ | 100.00% | 5228 | 100.00% | 6637 |
| | Rinott$_I$ | 99.32% | 5290 | 99.68% | 6725 |
| | ETSS$_C$ | 99.84% | 1288 | 99.96% | 1621 |
| | ETSS$_I$ | 94.22% | 1222 | 95.23% | 1542 |
| 30 | Rinott$_C$ | 100.00% | 4985 | 100.00% | 6287 |
| | Rinott$_I$ | 99.43% | 5017 | 99.79% | 6337 |
| | ETSS$_C$ | 99.95% | 1259 | 99.97% | 1549 |
| | ETSS$_I$ | 96.50% | 1250 | 97.36% | 1532 |

Table 2: Pre-Selection for Experiment 1

| | $P^* = 0.90$ | | $P^* = 0.95$ | |
|---|---|---|---|---|
| D | $n_0 = 20$ | $n_0 = 30$ | $n_0 = 20$ | $n_0 = 30$ |
| 1 | 9964/9970 | 9974/9977 | 9985/9969 | 9991/9984 |
| 2 | 9600/9718 | 9456/9606 | 9772/9731 | 9748/9654 |
| 3 | 8800/9116 | 8278/8663 | 9239/9102 | 8925/8671 |
| 4 | 7383/7992 | 6126/6744 | 8225/8034 | 7225/6821 |
| 5 | 5453/6324 | 3474/3267 | 6608/6309 | 4656/4334 |
| 6 | 3388/4373 | 1443/2067 | 4562/4322 | 2313/2081 |
| 7 | 1582/2413 | 373/642 | 2633/2502 | 781/699 |
| 8 | 686/1163 | 51/161 | 1176/1152 | 178/171 |
| 9 | 199/401 | 10/20 | 460/436 | 22/26 |
| 10 | 52/142 | 3/5 | 135/133 | 4/2 |

### 4.1 Experiment 1 Equal Variances

There are ten alternatives in the selection subset. Suppose $X_{ij} \sim \mathcal{N}(i, 6^2)$, $i = 1, 2, \ldots, 10$, where $\mathcal{N}(\mu, \sigma^2)$ denotes the normal distribution with mean $\mu$ and variance $\sigma^2$. We want to select an alternative with the minimum mean, i.e., alternative 1. The indifference amount $d^*$ is set to 0.90 for all cases. We estimate the achieved $P(CS)$ of Rinott$_I$ and Rinott$_C$ procedures as well as ETSS$_I$ and ETSS$_C$ procedures. To make these estimates we performed 10,000 independent experiments to compute $P(CS) =$ the proportion of experiments in which the best alternative (alternative 1) was selected, as well as $\overline{T}$ the average over the experiments of the total number of replications or batches ($T = \sum_{i=1}^{k} N_i$). We use two different initial replications, $n_0 = 20$ and 30.

The results of our experiment 1 are summarized in Table 1. Note that the $P(CS)$ are all larger than the specified $P^* = 0.90$ and 0.95. Rinott$_C$ and ETSS$_C$ have better coverage than Rinott$_I$ and ETSS$_I$ procedures with comparable total number of replications because CRNs are used. Moreover, the actual coverages of the Rinott$_C$ are 100%, because CRNs are completely synchronized when variances are equal, thus, the probability $P(\hat{\mu}_{i_1} < \hat{\mu}_{i_l}) = 1$, for $l = 2, 3, \ldots, k$ in this settings. Moreover, ETSS$_C$ gains more improvement in $P(CS)$ than Rinott$_C$.

Table 3: Pre-Selection for Experiment 1 using CRNs

| | $P^* = 0.90$ | | $P^* = 0.95$ | |
|---|---|---|---|---|
| D | $n_0 = 20$ | $n_0 = 30$ | $n_0 = 20$ | $n_0 = 30$ |
| 1 | 10K/10K | 10K/10K | 10K/10K | 10K/10K |
| 2 | 10K/1629 | 10K/636 | 10K/1676 | 10K/665 |
| 3 | 9999/539 | 9997/108 | 10K/555 | 10K/127 |
| 4 | 9874/70 | 9173/5 | 9970/78 | 9823/13 |
| 5 | 7903/0 | 2753/0 | 9287/0 | 5918/0 |
| 6 | 3315/0 | 124/0 | 5838/0 | 755/0 |
| 7 | 535/0 | 1/0 | 1969/0 | 27/0 |
| 8 | 37/0 | 0/0 | 294/0 | 2/0 |
| 9 | 1/0 | 0/0 | 22/0 | 0/0 |
| 10 | 0/0 | 0/0 | 1/0 | 0/0 |

Table 2 lists the results of subset pre-selection from both SSPS$_I$ and ETSS$_I$ procedures when samples are independent across alternatives. The $D$ column list the alternative designs. The $n_0$ columns list the frequencies of each alternative $i$ that are included in the subset for further simulation for both ETSS$_I$ and SSPS$_I$ with $P^* = 0.90$ and 0.95. For example, the first cell under $P^* = 0.90$ and $n_0 = 20$ indicates that 9964 out of those 10,000 experiments (i.e. 99.64%) alternative 1 has $N_i > n_0$ from ETSS and 9970 out of those 10,000 times (i.e. 99.70%) alternative 1 is included in the subset for further simulations from SSPS. The pre-selection performance of both procedures are about the same. Moreover, both procedures perform better with larger $n_0$.

Table 3 lists the pre-selection results when CRNs are used. The number $10K$ denotes 10,000. Both ETSS$_C$ and SSPS$_C$ perform better with CRNs. We are not able to show the validity of SSPS$_C$ (using CRNs with SSPS). However, empirically SSPS$_C$ performs much better than SSPS$_I$. In all cases, the subset contains no more than 4 alternatives. Equation (2) use correlated pairwise samples in different alternative to compute the variance, therefore, the variance is significantly reduced when CRNs are using. On the other hand, ETSS$_C$ uses correlated sample means to calculate $N_i$, consequently, the improvement is not as significant as SSPS$_C$.

### 4.2 Experiment 2 Increasing Variances

This is a variation of experiment 1. All settings are preserved except that the variance of each alternative increases as the mean increases: $X_{ij} \sim \mathcal{N}(i, (6 + i/2)^2)$, $i = 1, 2, \ldots, 10$.

The results are in Tables 4, 5 and 6. Because most alternatives have larger variance than experiment 1, more replications are needed since we are less confident of the best selection. Furthermore, the inferior alternatives are excluded

Table 4: P(CS) and Sample Sizes for Experiment 2

| | | $P^* = 0.90$ | | $P^* = 0.95$ | |
|---|---|---|---|---|---|
| $n_0$ | Procedure | $P(CS)$ | $\overline{T}$ | $P(CS)$ | $\overline{T}$ |
| 20 | $\text{Rinott}_C$ | 100.00% | 10182 | 100.00% | 12935 |
| | $\text{Rinott}_I$ | 99.43% | 10269 | 99.79% | 13023 |
| | $\text{ETSS}_C$ | 99.96% | 1582 | 99.98% | 2025 |
| | $\text{ETSS}_I$ | 93.14% | 1592 | 94.74% | 2038 |
| 30 | $\text{Rinott}_C$ | 100.00% | 9708 | 100.00% | 12239 |
| | $\text{Rinott}_I$ | 99.47% | 9761 | 99.81% | 12306 |
| | $\text{ETSS}_C$ | 100.00% | 1460 | 100.00% | 1825 |
| | $\text{ETSS}_I$ | 96.02% | 1528 | 96.89% | 1911 |

Table 5: Pre-Selection for Experiment 2

| | $P^* = 0.90$ | | $P^* = 0.95$ | |
|---|---|---|---|---|
| D | $n_0 = 20$ | $n_0 = 30$ | $n_0 = 20$ | $n_0 = 30$ |
| 1 | 9931/9958 | 9969/9973 | 9966/9953 | 9981/9972 |
| 2 | 9669/9773 | 9624/9682 | 9824/9802 | 9792/9691 |
| 3 | 9219/9345 | 8869/8975 | 9571/9355 | 9326/8988 |
| 4 | 8506/8631 | 7671/7808 | 9078/8614 | 8517/7874 |
| 5 | 7605/7699 | 6295/6358 | 8491/7728 | 7394/6407 |
| 6 | 6590/6588 | 4775/4700 | 7693/6706 | 5974/4675 |
| 7 | 5623/5533 | 3437/3286 | 6921/5633 | 4773/3362 |
| 8 | 4637/4457 | 2407/2242 | 5998/4442 | 3555/2218 |
| 9 | 3781/3580 | 1610/1413 | 5102/355 | 2584/1372 |
| 10 | 2970/2747 | 1021/814 | 4360/2750 | 1862/909 |

less often because the variances are larger. The $\text{Rinott}_C$ and $\text{ETSS}_C$ have better coverage than $\text{Rinott}_I$ and $\text{ETSS}_I$ procedures with comparable total simulation replications. The observed coverages of the $\text{Rinott}_C$ are 100%, even though the number of replications for each alternative are tremendously different. The pre-selection performance of $\text{ETSS}_I$ and $\text{SSPS}_I$ is about the same. However, $\text{SSPS}_C$ performs much better than $\text{ETSS}_C$. Even though inferior alternatives have larger variance, no more than 5 alternatives are included in the subset.

### 4.3 Experiment 3 decreasing Variances

This is another variation of experiment 1. All settings are preserved except that the variance of each alternative decreases as the mean increases: $X_{ij} \sim \mathcal{N}(i, (6 - i/2)^2)$, $i = 1, 2, \ldots, 10$.

The results are in Tables 7, 8 and 9. Because most alternatives have smaller variances than experiment 1, the total number of replications is smaller. The inferior alternatives are almost always excluded from the subset for further simulation because their variances are small. Once again, using CRNs improves $P(CS)$. The actual coverages of the $\text{Rinott}_C$ are 100%, even though the number of replications

Table 6: Pre-Selection for Experiment 2 using CRNs

| | $P^* = 0.90$ | | $P^* = 0.95$ | |
|---|---|---|---|---|
| D | $n_0 = 20$ | $n_0 = 30$ | $n_0 = 20$ | $n_0 = 30$ |
| 1 | 10K/10K | 10K/10K | 10K/10K | 10K/10K |
| 2 | 10K/2305 | 10K/1223 | 10K/2506 | 10K/1383 |
| 3 | 9999/826 | 10K/409 | 10K/908 | 10K/501 |
| 4 | 9975/258 | 9911/84 | 9997/300 | 9990/109 |
| 5 | 9753/8 | 8859/1 | 9935/8 | 9695/2 |
| 6 | 9016/0 | 6011/0 | 9664/0 | 8259/0 |
| 7 | 7561/0 | 3078/0 | 8985/0 | 5784/0 |
| 8 | 5856/0 | 1316/0 | 7839/0 | 3427/0 |
| 9 | 4204/0 | 535/0 | 6533/0 | 1816/0 |
| 10 | 2897/0 | 204/0 | 5202/0 | 880/0 |

Table 7: P(CS) and Sample Sizes for Experiment 3

| | | $P^* = 0.90$ | | $P^* = 0.95$ | |
|---|---|---|---|---|---|
| $n_0$ | Procedure | $P(CS)$ | $\overline{T}$ | $P(CS)$ | $\overline{T}$ |
| 20 | $\text{Rinott}_C$ | 100.00% | 2370 | 100.00% | 3009 |
| | $\text{Rinott}_I$ | 99.46% | 2391 | 99.75% | 3033 |
| | $\text{ETSS}_C$ | 99.64% | 1131 | 99.81% | 1407 |
| | $\text{ETSS}_I$ | 95.24% | 1047 | 96.07% | 1282 |
| 30 | $\text{Rinott}_C$ | 100.00% | 2255 | 100.00% | 2839 |
| | $\text{Rinott}_I$ | 99.34% | 2264 | 99.70% | 2855 |
| | $\text{ETSS}_C$ | 99.71% | 1138 | 99.84% | 1384 |
| | $\text{ETSS}_I$ | 96.66% | 1084 | 97.94% | 1320 |

are different for each alternative. In this setting, the pre-selection performance of $\text{ETSS}_I$ performs slightly better than $\text{SSPS}_I$. However, $\text{SSPS}_C$ has better performance than $\text{ETSS}_C$.

### 4.4 Experiment 4 Dependent Data Equal Variances

This is another variation of experiment 1 and is designed to check the robustness of these procedures to violation of the independence assumption. All settings are preserved except that data within each alternative are correlated. The input will be the *first-order autoregressive* (AR(1)) process, generated by the recurrence relation

$$X_{ij} = \mu_i + \varphi(X_{i,j-1} - \mu_i) + \epsilon_j, \quad \text{for} \quad j = 1, 2, \ldots,$$

where

$$E(\epsilon_j) = 0, \quad E(\epsilon_j \epsilon_k) = \begin{cases} \sigma^2 & \text{if } j = k \\ 0 & \text{otherwise} \end{cases},$$

$$0 < \varphi < 1,$$

Table 8: Pre-Selection for Experiment 3

| | $P^* = 0.90$ | | $P^* = 0.95$ | |
|---|---|---|---|---|
| D | $n_0 = 20$ | $n_0 = 30$ | $n_0 = 20$ | $n_0 = 30$ |
| 1 | 9963/9965 | 9983/9972 | 9991/9976 | 9995/9977 |
| 2 | 9532/9701 | 9374/9559 | 9736/9708 | 9640/9592 |
| 3 | 8255/8940 | 7746/8238 | 8791/8848 | 8183/8247 |
| 4 | 5394/7088 | 3558/5250 | 6398/7049 | 4686/5377 |
| 5 | 1979/4305 | 657/1929 | 2807/4177 | 1104/1929 |
| 6 | 232/1583 | 19/291 | 468/1553 | 49/286 |
| 7 | 7/296 | 0/11 | 11/258 | 0/13 |
| 8 | 0/26 | 0/0 | 0/18 | 0/0 |
| 9 | 0/0 | 0/0 | 0/0 | 0/0 |
| 10 | 0/0 | 0/0 | 0/0 | 0/0 |

Table 9: Pre-Selection for Experiment 3 using CRNs

| | $P^* = 0.90$ | | $P^* = 0.95$ | |
|---|---|---|---|---|
| D | $n_0 = 20$ | $n_0 = 30$ | $n_0 = 20$ | $n_0 = 30$ |
| 1 | 10K/10K | 10K/10K | 10K/10K | 10K/10K |
| 2 | 10K/1246 | 10K/491 | 10K/1339 | 10K/476 |
| 3 | 9990/355 | 9943/54 | 9994/365 | 9997/48 |
| 4 | 7608/14 | 3179/0 | 8982/16 | 5876/0 |
| 5 | 1126/0 | 25/0 | 2552/0 | 139/0 |
| 6 | 23/0 | 0/0 | 103/0 | 0/0 |
| 7 | 0/0 | 0/0 | 1/0 | 0/0 |
| 8 | 0/0 | 0/0 | 0/0 | 0/0 |
| 9 | 0/0 | 0/0 | 0/0 | 0/0 |
| 10 | 0/0 | 0/0 | 0/0 | 0/0 |

and $X_{i0}$ is a draw from the steady-state distribution. The AR(1) process shares many characteristics observed in simulation output processes, including first- and second-order stationarity, and autocorrelations that decline exponentially with increasing lag. If we make the additional assumption that the $\epsilon_j$'s are normally distributed, since we have already assumed that they are uncorrelated, they will now be independent as well, i.e., the $\epsilon_j$'s are i.i.d. $\mathcal{N}(0, \sigma^2)$. It can be shown that at $X_{ij}$ has a $\mathcal{N}(\mu_i, \frac{\sigma^2}{1-\varphi^2})$ distribution as $j \to \infty$. If we set $\mu_i = i$, $\varphi = 0.5$, and $\sigma^2 = 5^2$, then $X_{ij} \sim \mathcal{N}(i, 100/3)$, $i = 1, 2, \ldots, 10$.

The experimental results are in Tables 10, 11, and 12. All procedures allocate smaller sample sizes in this experiment than experiment 1, due to a smaller variance ($100/3 < 6^2$) of the $X_{ij}$ input sequence and perhaps dependence within in alternatives. These procedures underestimated the variance when data are not independent, consequently, the number of replications is not large enough and $P(CS)$ is not as good as in previous experiments. The degradation of the performance is less significant in Rinott$_I$ than ETSS$_I$ when data are not i.i.d. normal. However, using CRNs still improve $P(CS)$'s for both procedures. In fact,

Table 10: P(CS) and Sample Sizes for Experiment 4

| | | $P^* = 0.90$ | | $P^* = 0.95$ | |
|---|---|---|---|---|---|
| $n_0$ | Procedure | $P(CS)$ | $\overline{T}$ | $P(CS)$ | $\overline{T}$ |
| 20 | Rinott$_C$ | 99.86% | 4377 | 99.97% | 5555 |
| | Rinott$_I$ | 90.82% | 4430 | 93.71% | 5636 |
| | ETSS$_C$ | 93.60% | 1113 | 96.02% | 1399 |
| | ETSS$_I$ | 68.33% | 981 | 71.13% | 1229 |
| 30 | Rinott$_C$ | 99.99% | 4312 | 100.00% | 5435 |
| | Rinott$_I$ | 91.71% | 4347 | 93.65% | 5485 |
| | ETSS$_C$ | 94.90% | 1130 | 96.48% | 1385 |
| | ETSS$_I$ | 74.88% | 1054 | 77.28% | 1274 |

Table 11: Pre-Selection for Experiment 4

| | $P^* = 0.90$ | | $P^* = 0.95$ | |
|---|---|---|---|---|
| D | $n_0 = 20$ | $n_0 = 30$ | $n_0 = 20$ | $n_0 = 30$ |
| 1 | 10K/10K | 10K/9999 | 10K/9999 | 10K/10K |
| 2 | 9938/9951 | 9913/9913 | 9961/9938 | 9979/9916 |
| 3 | 9512/9533 | 9217/9357 | 9771/9631 | 9586/9288 |
| 4 | 8170/8628 | 6593/7234 | 8975/8599 | 7914/7231 |
| 5 | 5666/6684 | 2861/3796 | 7149/6616 | 4388/3817 |
| 6 | 2874/4023 | 617/1242 | 4373/4040 | 1385/1234 |
| 7 | 1080/1940 | 76/234 | 1954/1942 | 217/221 |
| 8 | 229/679 | 4/23 | 636/691 | 23/21 |
| 9 | 71/189 | 0/0 | 147/198 | 2/0 |
| 10 | 2/41 | 0/0 | 37/58 | 0/0 |

the observed $P(CS)$'s from the Rinott$_C$ and ETSS$_C$ are all higher than the specified $P^*$ despite the autocorrelation in the input data. The performance of subset pre-selection is as good as experiment 1 because with CRNs the difference of the sample means was not significantly distorted, which preserve the performance of pre-selection even when data are correlated. Both SSPS$_C$ and ETSS$_C$ procedures obtain significant improvement of $P(CS)$ with CRNs in this settings.

## 5 CONCLUSIONS

Our experimental results show that the marginal effort required for using CRNs is minimal, yet the achieved efficiency improvement is significant. We point out that when using CRNs we can benefit from $P_C(E_2) \geq P_I(E_2)$ and maybe $P_C(E_l|E_2, E_3, \ldots, E_{l-1}) \geq P_I(E_l|E_2, E_3, \ldots, E_{l-1})$, for $l = 3, 4, \ldots, k$. Our experimental results also show that the intrinsic subset pre-selection of the ETSS$_I$ procedure performs as well as the SSPS$_I$ in normal cases, i.e., independent sampling across alternatives. Moreover, the ETSS procedure is easier to implement and requires less computation than SSPS.

Table 12: Pre-Selection for Experiment 4 using CRNS

| | $P^* = 0.90$ | | $P^* = 0.95$ | |
|---|---|---|---|---|
| D | $n_0 = 20$ | $n_0 = 30$ | $n_0 = 20$ | $n_0 = 30$ |
| 1 | 10K/10K | 10K/10K | 10K/10K | 10K/10K |
| 2 | 10K/982 | 10K/263 | 10K/1065 | 10K/261 |
| 3 | 9993/244 | 9986/24 | 9999/221 | 10K/18 |
| 4 | 9409/43 | 7742/3 | 9789/38 | 9267/3 |
| 5 | 6092/5 | 1829/1 | 7963/7 | 4134/0 |
| 6 | 2213/0 | 97/0 | 4214/1 | 610/0 |
| 7 | 427/0 | 2/0 | 1391/0 | 30/0 |
| 8 | 46/0 | 0/0 | 281/0 | 1/0 |
| 9 | 5/0 | 0/0 | 32/0 | 0/0 |
| 10 | 0/0 | 0/0 | 4/0 | 0/0 |

The average sample size allocated by the $\text{ETSS}_C$ is generally larger than $\text{ETSS}_I$ in our experiments. This is because using CRNs improve the performance of subset pre-selection. The procedure correctly increases the number of promising alternatives being selected in the subset. That is while the reduction of variance reduce the sample size, the increase in the precision of the difference of the sample size increase the sample size. Thus, the final sample size general increases with CRNs.

## ACKNOWLEDGMENTS

I would like to thank Professors W. David Kelton and Barry L. Nelson for their assistance, and the anonymous referee for his comments.

## Appendix A. Proof of Proposition I

**Proof:** Let $F_N(x)$ denote the cumulative density function of the standard normal distribution. If $x_1$ has a $\mathcal{N}(\mu, \sigma_1^2)$ and $x_2$ has a $\mathcal{N}(\mu, \sigma_2^2)$ distribution and $\sigma_1 > \sigma_2$, then for any $\mu > 0 \geq y$, $F_N(\frac{y-\mu}{\sigma_1}) > F_N(\frac{y-\mu}{\sigma_2})$; because $0 > \frac{y-\mu}{\sigma_1} > \frac{y-\mu}{\sigma_2}$ and $F_N(x)$ is a monotonically increasing function.

Let $a = \mu_{i_l} - \mu_{i_1}$, then $a > 0$. Since $X_{ij}$'s are i.i.d. normal within each alternative, $\hat{\mu}_{i_l} - \hat{\mu}_{i_1} = \hat{a}$ also has a normal distribution. Therefore, we will have a smaller probability of $\hat{a} < 0$, i.e., the probability of incorrect selection, with a smaller variance of $\hat{a}$. Let $\sigma_C$ denote the variance of $\hat{a}$ when common random numbers are used, and let $\sigma_I$ denote the variance of $\hat{a}$ with independent sampling. From equation (7), we know that $\sigma_I \geq \sigma_C$. Therefore, $P_C(CS) \geq P_I(CS)$ ∎

## REFERENCES

Bechhofer, R. E., T. J. Santner and D. Goldsman. 1995. *Design and Analysis of Experiments for Statistical Selection, Screening and Multiple Comparisons*. New York: John Wiley & Sons, Inc.

Chen, E. J., and W. D. Kelton. 2000a. Mean estimation based on $\phi$-mixing sequences. In *Proceedings of the* $33^{rd}$ *Annual Simulation Symposium*, ed. D.C. Young, 237–244. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Chen, E. J., and W. D. Kelton. 2000b. A Stopping Procedure Based on Phi-Mixing Conditions. In *Proceedings of the 2000 Winter Simulation Conference*, ed. J.A. Joines, R. Barton, P. Fishwick, and K. Kang, 617–626. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Chen, E. J., and W. D. Kelton. 2000c. An Enhanced Two-Stage Selection Procedure. In *Proceedings of the 2000 Winter Simulation Conference*, ed. J.A. Joines, R. Barton, P. Fishwick, and K. Kang, 727–735. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Chick, S. E., and K. Inoue. 2000. New Results on Procedures that select the best system using CRN. In *Proceedings of the 2000 Winter Simulation Conference*, ed. J.A. Joines, R. Barton, P. Fishwick, and K. Kang, 554–561. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Clark, G. M., and W. N. Yang. 1986. A Bonferroni Selection Procedure When Using Common Random Numbers with Unknown Variances. In *Proceedings of the 1986 Winter Simulation Conference*, ed. J.R. Wilson, J.O. Henriksen, and S.D. Roberts, 313–315. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Dudewicz, E. J. and S. R. Dalal. 1975. Allocation of Observations in Ranking and Selection with Unequal Variances. *Sankhya* B37: 28–78.

Goldsman D., and B. L. Nelson. 1998. Statistical Screening, Selection, and Multiple Comparison Procedures in Computer Simulation. In *Proceedings of the 1998 Winter Simulation Conference*, ed. D.J. Medeiros, E.F Watson, J.S. Carson, and M.S. Manivannan, 159–166. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Goldsman D., B. L. Nelson, T. Opicka, and A. A. B. Pritsker. 1999. A Ranking and Selection Project: Experiences from A University-Industry Collaboration. In *Proceedings of the 1999 Winter Simulation Conference*, ed. P.A. Farrington, H.B. Nembhard, D.T. Sturrock, and G.W. Evans, 83–92. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Law, A. M., and W. D. Kelton. 2000. *Simulation Modeling and Analysis*. 3rd ed. New York: McGraw-Hill.

Nelson, B. L. and F. J. Matejcik. 1995. Using Common Random Numbers for Indifference-zone Selection

and Multiple Comparisons in Simulation. *Management Science* 41: 1935–1945.

Rinott, Y., 1978. On Two-stage Selection Procedures and Related Probability Inequalities. *Communications in Statistics* A7: 799-811.

Tong, Y. L. 1980. *Probability Inequalities in Multivariate Distributions.* New York: Academic Press.

Wilcox, R. R., 1984. A Table for Rinott's Selection Procedure. *Journal of Quality Technology* Vol. 16, No. 2: 97-100.

Wright, R. D. and T. E. Ramsay, Jr., 1979. On the Effectiveness of Common Random Numbers. *Management Science* 25: 649–656.

## AUTHOR BIOGRAPHY

**E. JACK CHEN** is a Senior Staff Specialist with BASF Corporation. He received a Ph.D. from University of Cincinnati. His research interests are in the area of computer simulation. His email and web addresses are <chenej@basf.com> and <www.econqa.cba.uc.edu/~chenj>.